

SEQUENCE LISTING

<110> DAVIS, BENJAMIN GUY

<120> MODIFIED CARBOHYDRATE PROCESSING ENZYME

<130> 117-480 / N84207A PJC

<140> US 10/

<141> 2003-10-07

<150> US 60/416,263

<151> 2002-10-07

<160> 23

<170> PatentIn version 3.1

<210> 1

<211> 2346

<212> DNA

<213> Sulfolobus solfataricus

<220>

<221> CDS

<222> (229)..(1698)

<223>

<400> 1

aaggagaaac ttggcagttt ataacttgac agtaggttgt ggagtgatga ctggatcaat 60

actaggagga gtagcatata attacgttac acaattttat aaccaatat attcaataga 120

ccttatgctt atcctatcct ctattctaag attctcggtt tctcccctat tcttgaccat 180

aaaagatact cgctcaaagc ttaaataata ttaatcataa ataaagtc atg tac tca 237
Met Tyr Ser
1

ttt cca aat agc ttt agg ttt ggt tgg tcc cag gcc gga ttt caa tca 285
Phe Pro Asn Ser Phe Arg Phe Gly Trp Ser Gln Ala Gly Phe Gln Ser
5 10 15

gaa atg gga aca cca ggg tca gaa gat cca aat act gac tgg tat aaa 333
Glu Met Gly Thr Pro Gly Ser Glu Asp Pro Asn Thr Asp Trp Tyr Lys
20 25 30 35

tgg gtt cat gat cca gaa aac atg gca gcg gga tta gta agt gga gat 381
Trp Val His Asp Pro Glu Asn Met Ala Ala Gly Leu Val Ser Gly Asp
40 45 50

cta cca gaa aat ggg cca ggc tac tgg gga aac tat aag aca ttt cac 429
Leu Pro Glu Asn Gly Pro Gly Tyr Trp Gly Asn Tyr Lys Thr Phe His
55 60 65

gat aat gca caa aaa atg gga tta aaa ata gct aga cta aat gtg gaa Asp Asn Ala Gln Lys Met Gly Leu Lys Ile Ala Arg Leu Asn Val Glu 70 75 80	477
tgg tct agg ata ttt cct aat cca tta cca agg cca caa aac ttt gat Trp Ser Arg Ile Phe Pro Asn Pro Leu Pro Arg Pro Gln Asn Phe Asp 85 90 95	525
gaa tca aaa caa gat gtg aca gag gtt gag ata aac gaa aac gag tta Glu Ser Lys Gln Asp Val Thr Glu Val Glu Ile Asn Glu Asn Glu Leu 100 105 110 115	573
aag aga ctt gac gag tac gct aat aaa gac gca tta aac cat tac agg Lys Arg Leu Asp Glu Tyr Ala Asn Lys Asp Ala Leu Asn His Tyr Arg 120 125 130	621
gaa ata ttc aag gat ctt aaa agt aga gga ctt tac ttt ata cta aac Glu Ile Phe Lys Asp Leu Lys Ser Arg Gly Leu Tyr Phe Ile Leu Asn 135 140 145	669
atg tat cat tgg cca tta cct cta tgg tta cac gac cca ata aga gta Met Tyr His Trp Pro Leu Pro Leu Trp Leu His Asp Pro Ile Arg Val 150 155 160	717
aga aga gga gat ttt act gga cca agt ggt tgg cta agt act aga aca Arg Arg Gly Asp Phe Thr Gly Pro Ser Gly Trp Leu Ser Thr Arg Thr 165 170 175	765
gtt tac gaa ttc gct aga ttc tca gct tat ata gct tgg aaa ttc gat Val Tyr Glu Phe Ala Arg Phe Ser Ala Tyr Ile Ala Trp Lys Phe Asp 180 185 190 195	813
gat cta gtg gat gag tac tca aca atg aat gaa cct aac gtt gtt gga Asp Leu Val Asp Glu Tyr Ser Thr Met Asn Glu Pro Asn Val Val Gly 200 205 210	861
ggt tta gga tac gtt ggt gtt aag tcc ggt ttt ccc cca gga tac cta Gly Leu Gly Tyr Val Gly Val Lys Ser Gly Phe Pro Pro Gly Tyr Leu 215 220 225	909
agc ttt gaa ctt tcc cgt agg cat atg tat aac atc att caa gct cac Ser Phe Glu Leu Ser Arg Arg His Met Tyr Asn Ile Ile Gln Ala His 230 235 240	957
gca aga gcg tat gat ggg ata aag agt gtt tct aaa aaa cca gtt gga Ala Arg Ala Tyr Asp Gly Ile Lys Ser Val Ser Lys Lys Pro Val Gly 245 250 255	1005
att att tac gct aat agc tca ttc cag ccg tta acg gat aaa gat atg Ile Ile Tyr Ala Asn Ser Ser Phe Gln Pro Leu Thr Asp Lys Asp Met 260 265 270 275	1053
gaa gcg gta gag atg gct gaa aat gat aat aga tgg tgg ttc ttt gat Glu Ala Val Glu Met Ala Glu Asn Asp Asn Arg Trp Trp Phe Phe Asp 280 285 290	1101

gct ata ata aga ggt gag atc acc aga gga aac gag aag att gta aga	1149
Ala Ile Ile Arg Gly Glu Ile Thr Arg Gly Asn Glu Lys Ile Val Arg	
295 300 305	
gat gac cta aag ggt aga ttg gat tgg att gga gtt aat tat tac act	1197
Asp Asp Leu Lys Gly Arg Leu Asp Trp Ile Gly Val Asn Tyr Tyr Thr	
310 315 320	
agg act gtt gtg aag agg act gaa aag gga tac gtt agc tta gga ggt	1245
Arg Thr Val Val Lys Arg Thr Glu Lys Gly Tyr Val Ser Leu Gly Gly	
325 330 335	
tac ggt cac gga tgt gag agg aat tct gta agt tta gcg gga tta cca	1293
Tyr Gly His Gly Cys Glu Arg Asn Ser Val Ser Leu Ala Gly Leu Pro	
340 345 350 355	
acc agc gac ttc ggc tgg gag ttc ttc cca gaa ggt tta tat gac gtt	1341
Thr Ser Asp Phe Gly Trp Glu Phe Phe Pro Glu Gly Leu Tyr Asp Val	
360 365 370	
ttg acg aaa tac tgg aat aga tat cat ctc tat atg tac gtt act gaa	1389
Leu Thr Lys Tyr Trp Asn Arg Tyr His Leu Tyr Met Tyr Val Thr Glu	
375 380 385	
aat ggt att gcg gat gat gcc gat tat caa agg ccc tat tat tta gta	1437
Asn Gly Ile Ala Asp Asp Ala Asp Tyr Gln Arg Pro Tyr Tyr Leu Val	
390 395 400	
tct cac gtt tat caa gtt cat aga gca ata aat agt ggt gca gat gtt	1485
Ser His Val Tyr Gln Val His Arg Ala Ile Asn Ser Gly Ala Asp Val	
405 410 415	
aga ggg tat tta cat tgg tct cta gct gat aat tac gaa tgg gct tca	1533
Arg Gly Tyr Leu His Trp Ser Leu Ala Asp Asn Tyr Glu Trp Ala Ser	
420 425 430 435	
gga ttc tct atg agg ttt ggt ctg tta aag gtc gat tac aac act aag	1581
Gly Phe Ser Met Arg Phe Gly Leu Leu Lys Val Asp Tyr Asn Thr Lys	
440 445 450	
aga cta tac tgg aga ccc tca gca cta gta tat agg gaa atc gcc aca	1629
Arg Leu Tyr Trp Arg Pro Ser Ala Leu Val Tyr Arg Glu Ile Ala Thr	
455 460 465	
aat ggc gca ata act gat gaa ata gag cac tta aat agc gta cct cca	1677
Asn Gly Ala Ile Thr Asp Glu Ile Glu His Leu Asn Ser Val Pro Pro	
470 475 480	
gta aag cca tta agg cac taa actttctcaa gtctcactat accaaatgag	1728
Val Lys Pro Leu Arg His	
485	
ttttctttta atcttattct aatctcattt tcattagatt gcaatacttt cataccttct	1788
atattattta ttttgtacct tttgggatct acacttaatg ttagcctaata ttgaaagtca	1848
tttagattta atactgttac cagtccatcc cttttaatta ttaatgaaaa taagaaggga	1908

taagtagcga tagcccttat tccgatatgg tctccaacaa tatcccttat tatctgcctt 1968
 gcaacactag ggtagaactc tgaaatcaga tatggtaggt aagttgtaag tgataggacg 2028
 taaacttttag agttagagta agtgttctga aagactactg ggtgcaattc gacaccgtta 2088
 taggcgtaaa ggattggcgt agctccgttt aatgaaaata taggtcctac agggaaattg 2148
 gcttgccctct tgtaatatga ccaatagaac gttttcccat ccctgggttaa cgcattgaca 2208
 ctaacactat cgtaaataca gttaccgaca ccaagaattt tcagtgcagt atcccccaag 2268
 acttcaataa gcttttttagc tgcacttgct gtaaacatta agttaactcc cctattaagt 2328
 aaatccacaa tatctaga 2346

<210> 2
 <211> 489
 <212> PRT
 <213> Sulfolobus solfataricus

<400> 2
 Met Tyr Ser Phe Pro Asn Ser Phe Arg Phe Gly Trp Ser Gln Ala Gly
 1 5 10 15
 Phe Gln Ser Glu Met Gly Thr Pro Gly Ser Glu Asp Pro Asn Thr Asp
 20 25 30
 Trp Tyr Lys Trp Val His Asp Pro Glu Asn Met Ala Ala Gly Leu Val
 35 40 45
 Ser Gly Asp Leu Pro Glu Asn Gly Pro Gly Tyr Trp Gly Asn Tyr Lys
 50 55 60
 Thr Phe His Asp Asn Ala Gln Lys Met Gly Leu Lys Ile Ala Arg Leu
 65 70 75 80
 Asn Val Glu Trp Ser Arg Ile Phe Pro Asn Pro Leu Pro Arg Pro Gln
 85 90 95
 Asn Phe Asp Glu Ser Lys Gln Asp Val Thr Glu Val Glu Ile Asn Glu
 100 105 110
 Asn Glu Leu Lys Arg Leu Asp Glu Tyr Ala Asn Lys Asp Ala Leu Asn
 115 120 125
 His Tyr Arg Glu Ile Phe Lys Asp Leu Lys Ser Arg Gly Leu Tyr Phe
 130 135 140
 Ile Leu Asn Met Tyr His Trp Pro Leu Pro Leu Trp Leu His Asp Pro
 145 150 155 160
 Ile Arg Val Arg Arg Gly Asp Phe Thr Gly Pro Ser Gly Trp Leu Ser
 165 170 175

Thr	Arg	Thr	Val	Tyr	Glu	Phe	Ala	Arg	Phe	Ser	Ala	Tyr	Ile	Ala	Trp	180	185	190	
Lys	Phe	Asp	Asp	Leu	Val	Asp	Glu	Tyr	Ser	Thr	Met	Asn	Glu	Pro	Asn	195	200	205	
Val	Val	Gly	Gly	Leu	Gly	Tyr	Val	Gly	Val	Lys	Ser	Gly	Phe	Pro	Pro	210	215	220	
Gly	Tyr	Leu	Ser	Phe	Glu	Leu	Ser	Arg	Arg	His	Met	Tyr	Asn	Ile	Ile	225	230	235	240
Gln	Ala	His	Ala	Arg	Ala	Tyr	Asp	Gly	Ile	Lys	Ser	Val	Ser	Lys	Lys	245	250	255	
Pro	Val	Gly	Ile	Ile	Tyr	Ala	Asn	Ser	Ser	Phe	Gln	Pro	Leu	Thr	Asp	260	265	270	
Lys	Asp	Met	Glu	Ala	Val	Glu	Met	Ala	Glu	Asn	Asp	Asn	Arg	Trp	Trp	275	280	285	
Phe	Phe	Asp	Ala	Ile	Ile	Arg	Gly	Glu	Ile	Thr	Arg	Gly	Asn	Glu	Lys	290	295	300	
Ile	Val	Arg	Asp	Asp	Leu	Lys	Gly	Arg	Leu	Asp	Trp	Ile	Gly	Val	Asn	305	310	315	320
Tyr	Tyr	Thr	Arg	Thr	Val	Val	Lys	Arg	Thr	Glu	Lys	Gly	Tyr	Val	Ser	325	330	335	
Leu	Gly	Gly	Tyr	Gly	His	Gly	Cys	Glu	Arg	Asn	Ser	Val	Ser	Leu	Ala	340	345	350	
Gly	Leu	Pro	Thr	Ser	Asp	Phe	Gly	Trp	Glu	Phe	Phe	Pro	Glu	Gly	Leu	355	360	365	
Tyr	Asp	Val	Leu	Thr	Lys	Tyr	Trp	Asn	Arg	Tyr	His	Leu	Tyr	Met	Tyr	370	375	380	
Val	Thr	Glu	Asn	Gly	Ile	Ala	Asp	Asp	Ala	Asp	Tyr	Gln	Arg	Pro	Tyr	385	390	395	400
Tyr	Leu	Val	Ser	His	Val	Tyr	Gln	Val	His	Arg	Ala	Ile	Asn	Ser	Gly	405	410	415	
Ala	Asp	Val	Arg	Gly	Tyr	Leu	His	Trp	Ser	Leu	Ala	Asp	Asn	Tyr	Glu	420	425	430	
Trp	Ala	Ser	Gly	Phe	Ser	Met	Arg	Phe	Gly	Leu	Leu	Lys	Val	Asp	Tyr	435	440	445	
Asn	Thr	Lys	Arg	Leu	Tyr	Trp	Arg	Pro	Ser	Ala	Leu	Val	Tyr	Arg	Glu	450	455	460	
Ile	Ala	Thr	Asn	Gly	Ala	Ile	Thr	Asp	Glu	Ile	Glu	His	Leu	Asn	Ser	465	470	475	480

Val Pro Pro Val Lys Pro Leu Arg His
485

<210> 3
<211> 25
<212> PRT
<213> Sulfolobus shibatae

<400> 3
Trp Ser Leu Ala Asp Asn Tyr Glu Trp Ala Ser Gly Phe Ser Met Arg
1 5 10 15

Phe Gly Leu Leu Lys Val Asp Tyr Asn
20 25

<210> 4
<211> 25
<212> PRT
<213> Sulfolobus acidocaldarius

<400> 4
Trp Ser Leu Ala Asp Asn Tyr Glu Trp Ser Ser Gly Phe Ser Met Arg
1 5 10 15

Phe Gly Leu Leu Lys Val Asp Tyr Leu
20 25

<210> 5
<211> 25
<212> PRT
<213> Thermoplasma volcanium

<400> 5
Trp Ser Leu Ala Asp Asn Tyr Glu Trp Ala Ser Gly Phe Ser Met Arg
1 5 10 15

Phe Gly Leu Leu Lys Val Asp Tyr Asn
20 25

<210> 6
<211> 25
<212> PRT
<213> Pyrococcus furiosus

<400> 6
Trp Ser Leu Thr Asp Asn Tyr Glu Trp Ala Gln Gly Phe Arg Met Arg
1 5 10 15

Phe Gly Leu Val Tyr Val Asp Phe Glu
20 25

<210> 7

<211> 25
<212> PRT
<213> Agrobacterium tumefaciens

<400> 7
Trp Ser Leu Met Asp Asn Phe Glu Trp Ala Glu Gly Tyr Arg Met Arg
1 5 10 15

Phe Gly Leu Val His Val Asp Tyr Glu
20 25

<210> 8
<211> 25
<212> PRT
<213> Bacillus circulans

<400> 8
Trp Ser Leu Met Asp Asn Phe Glu Trp Ala Glu Gly Tyr Gly Met Arg
1 5 10 15

Phe Gly Leu Val His Val Asp Tyr Asp
20 25

<210> 9
<211> 25
<212> PRT
<213> Agrobacterium sp. (strain ATCC 21400)

<400> 9
Trp Ser Leu Met Asp Asn Phe Glu Trp Ala Glu Gly Tyr Arg Met Arg
1 5 10 15

Phe Gly Leu Val His Val Asp Tyr Gln
20 25

<210> 10
<211> 25
<212> PRT
<213> Rhizobium meliloti

<400> 10
Trp Ser Leu Met Asp Asn Phe Glu Trp Ala Glu Gly Tyr Arg Met Arg
1 5 10 15

Phe Gly Ile Val His Val Asp Tyr Glu
20 25

<210> 11
<211> 25
<212> PRT
<213> Bacillus halodurans

<400> 11

Trp Ser Leu Leu Asp Asn Phe Glu Trp Ala Glu Gly Tyr Ser Met Arg
 1 5 10 15

Phe Gly Ile Val His Val Asn Tyr Arg
 20 25

<210> 12
 <211> 25
 <212> PRT
 <213> *Paenibacillus polymyxa*

<400> 12
 Trp Ser Leu Leu Asp Asn Phe Glu Trp Ala Glu Gly Tyr Asn Met Arg
 1 5 10 15

Phe Gly Met Ile His Val Asp Phe Arg
 20 25

<210> 13
 <211> 25
 <212> PRT
 <213> *Pyrococcus woesi*

<400> 13
 Trp Ser Leu Ala Asp Asn Tyr Glu Trp Ala Ser Gly Phe Ser Met Arg
 1 5 10 15

Phe Gly Leu Leu His Val Asp Tyr Asn
 20 25

<210> 14
 <211> 25
 <212> PRT
 <213> *Dalbergia cochinchinensis*

<400> 14
 Trp Ser Leu Leu Asp Asn Phe Glu Trp Ala Glu Gly Tyr Thr Ser Arg
 1 5 10 15

Phe Gly Leu Tyr Phe Val Asn Tyr Thr
 20 25

<210> 15
 <211> 25
 <212> PRT
 <213> *Costus speciosus*

<400> 15
 Trp Ala Leu Thr Asp Asn Phe Glu Trp Asp Lys Gly Tyr Thr Glu Arg
 1 5 10 15

Phe Gly Leu Ile Tyr Ile Asp Tyr Asp
 20 25

<210> 16
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 16
 Arg Ser Leu Ile Asp Gly Phe Glu Gly Pro Ser Gly Tyr Ser Gln Arg
 1 5 10 15
 Phe Gly Leu His His Val Asn Phe Ser
 20 25

<210> 17
 <211> 25
 <212> PRT
 <213> Sinapis alba

<400> 17
 Trp Ala Leu Gly Asp Asn Tyr Glu Phe Cys Lys Gly Phe Thr Val Arg
 1 5 10 15
 Phe Gly Leu Ser Tyr Val Asn Trp Asp
 20 25

<210> 18
 <211> 25
 <212> PRT
 <213> Staphylococcus aureus

<400> 18
 Trp Ser Leu Met Asp Val Phe Ser Trp Ser Asn Gly Tyr Glu Lys Arg
 1 5 10 15
 Tyr Gly Leu Phe Tyr Val Asp Phe Glu
 20 25

<210> 19
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide primer

<400> 19
 ccatgggaca ccaccaccac caccaccact cattac

36

<210> 20
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide primer

 <400> 20
 ctcgagtttag tgcctttatg gctttactgg aggtac 36

<210> 21
 <211> 35
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide primer

 <400> 21
 tctagctgat aattactggt gggcttcagg attct 35

<210> 22
 <211> 35
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide primer

 <400> 22
 ctagctgata attacgaatg tgcttcagga ttctc 35

<210> 23
 <211> 30
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide primer

 <400> 23
 gcttcaggat tctctttag gtttggtctg 30